Genetic characterization of *Vibrio cholerae* strains by inter simple sequence repeat-PCR

A. Ravi Kumar¹, V. Sathish¹, G. Balakrish Nair² & J. Nagaraju¹

¹Laboratory of Molecular Genetics, Centre for DNA Fingerprinting and Diagnostics, Hyderabad, India; and ²Laboratory Sciences Division, ICDDR, B Centre for Health and Population Research, Mohakhali, Dhaka, Bangladesh

**Correspondence:** J. Nagaraju, Laboratory of Molecular Genetics, Centre for DNA Fingerprinting and Diagnostics, ECIL Road, Nacharam, Hyderabad 500076 India. Tel.: +91 40 2715 1344; fax: +91 40 2715 5610; email: jnagaraju@cdfd.org.in.

Received 20 February 2007; revised 14 April 2007; accepted 14 April 2007.

DOI:10.1111/j.1574-6968.2007.00762.x

**Editor:** Marco Soria

**Keywords:** *Vibrio cholerae*; inter simple sequence repeat-PCR; phylogeny.

**Abstract**

The utility of inter simple sequence repeat-PCR (ISSR-PCR) assay in the characterization and elucidation of the phylogenetic relationship between the pathogenic and nonpathogenic isolates of *Vibrio cholerae* is demonstrated. A total of 45 *V. cholerae* strains including 15 O1 *El Tor*, nine O139 and 21 non-O1/non-O139 strains were analyzed using eight ISSR primers. These primers, which are essentially simple sequence repeats (SSR) with additional nonrepeat bases at the 5’ or 3’ end, amplify genomic regions interspersed between closely spaced SSRs. Neighbor-joining analysis showed that the strains belonging to the same serogroup clustered together with the exception of one O1 and two O139 strains. The absence of pathogenicity islands in these strains, as confirmed by PCR, suggested their non-O1/non-O139 origin. Thus the ISSR-PCR-based phylogeny was consistent with the classification of *V. cholerae* based on serological methods. A finer resolution of the clustering of the toxinogenic O1 *El Tor* and toxinogenic O139 subtypes was obtained by ISSR-PCR analysis as compared with the Enterobacterial Repetitive Intergenic Consensus sequences-based PCR analysis for the same set of strains. Thus, it is proposed that ISSR-PCR is an efficient tool in phylogenetic classification of prokaryotic genomes in general and diagnostic genotyping of microbial pathogens in particular.

**Introduction**

The gram-negative bacterium *Vibrio cholerae* has been identified as the main causative agent of the acute dehydrating diarrhoeal condition called cholera. Among more than 200 O-antigen serogroups that have been identified and characterized, only two serogroups -O1 and O139 – are known to cause the disease, which can be epidemic, endemic or pandemic in nature (Shimada *et al.*, 1994). The strains belonging to the O1 serogroup are further classified into two biotypes, namely the Classical and *El Tor* biotypes, which can be differentiated by biochemical and genetic traits (Kaper *et al.*, 1995). Other *V. cholerae* serogroups, which are not associated with epidemics or pandemics, are collectively referred to as the non-O1/non-O139 serogroup (Albert & Nair, 2005). The toxinogenic O1 and O139 strains contain the Vibrio Pathogenicity Island (VPI) and CTXΦ filamentous bacteriophages, but 95% of non-O1/non-O139 *V. cholerae* strains do not harbour these elements. The toxinogenic strains are believed to have evolved from ancestral nontoxinogenic *V. cholerae* strains upon acquisition of the VPI and CTXΦ filamentous bacteriophages by horizontal gene transfer (Hacker *et al.*, 1997). Several studies have shown that O139 strains are phylogenetically and serotypically very similar to the O1 *El Tor* strains (Ramamurthy *et al.*, 1993; Bik *et al.*, 1995; Stroeher *et al.*, 1995; Dalsgaard *et al.*, 1998), supporting the event of horizontal gene transfer between the O1 and nontoxinogenic O139 serogroups. Reports on non-O1/non-O139 strains as the causative agents of sporadic cases of a cholera-like disease (Morris, 1994) and isolated outbreaks (Bagchi *et al.*, 1993) also point towards the emergence of newer variants of toxinogenic *V. cholerae* from non-O1/non-O139 serogroups.

In view of the propensity of O1 and O139 to cause global epidemics and pandemics, an understanding of the genetic structure of microbial populations is relevant to track the source and spread of epidemics, to manage the spread of resistance to drugs and to improve the efficacy and safety of genetically engineered microorganisms intended for environmental applications.

Unlike biochemical and serological methods, molecular tools for genotyping of strains are more accurate in terms of
identifying sublineages of microorganism that would enhance the sensitivity of epidemiological investigations. Various molecular techniques such as pulse field gel electrophoresis (PFGE) (Safa et al., 2005), randomly amplified polymorphic DNA (RAPD) (Makino et al., 1995), amplified fragment length polymorphism (AFLP) (Jiang et al., 2000), repetitive extragenic palindromic PCR (REP-PCR) (Shangkuan et al., 1997), Enterobacterial Repetitive Intergenic Consensus sequences-based PCR (ERIC-PCR) (Versalovic et al., 1991), ribotyping (Bhanumathi et al., 2002), Multi Locus Sequence Typing (MLST) (Lee et al., 2006) and, finally, comparative genomics (Beres et al., 2006) have been applied in attempts to characterize and establish the genetic population structure of microorganisms. Difference in sensitivity and accuracy of different methods reflects upon the need for more efficient genotyping techniques in the microbiologists’ tool kit.

The presence of simple sequence repeats (SSR) in prokaryotes is well documented (Gur-Arie et al., 2000), and some SSRs show extensive length polymorphisms (Yang et al., 2003; Sreenu et al., 2006). Successful use of PCR-based SSR amplification followed by amplicon size determination to analyze the spread of microbial pathogens has been reported for Haemophilus influenzae and Candida albicans (Bretagne et al., 1997; van Belkum et al., 1997). Inter simple sequence repeat-PCR (ISSR-PCR) technique also exploits the genome-wide distribution of SSRs. Single oligonucleotide primers containing a frequently occurring SSR motif with a stretch of arbitrary nucleotides, anchored at either the 5’ or 3’ end, initiate PCR amplification of genomic segments flanked by inversely oriented, closely spaced repeats (Fig. 1). The PCR products thus generated reveal multilocus profiles, which could be revealed on agarose or polyacrylamide gels (Zietkiewicz et al., 1994). ISSR-PCR analysis is technically simple and requires no prior genomic sequence information. The method yields highly reproducible results and generates abundant polymorphisms in eukaryotic systems. To date, many studies have been reported on using ISSR markers for identification or genotyping of variants in agriculturally important plants like wheat, rice and recently in insects (Nagaoka & Ogihara, 1997; Reddy et al., 1999; Nagaraju et al., 2002), but until now the feasibility of employing ISSR markers for diagnostic purpose or for fingerprinting of microbial pathogens has not been addressed.

In the present study, the utility of ISSR-PCR as a novel method of genotyping of V. cholerae has been demonstrated. The results show that as a diagnostic tool, ISSR-PCR could be invaluable in disease diagnosis, recognizing outbreaks of infection, and detecting the source and spread of infection.

Materials and methods

SSR analysis

The whole genome sequence of V. cholerae El Tor strain N16961 was analyzed for the abundance and distribution of SSRs using the prokaryotic microsatellite database, MICdb (http://210.212.212.7/MIC/index.html) (Sreenu et al., 2006). The whole genome sequence of V. cholerae El Tor strain N16961 was analyzed for the abundance and distribution of SSRs using the prokaryotic microsatellite database, MICdb (http://210.212.212.7/MIC/index.html) (Sreenu et al., 2006).
Preparation of genomic DNA

Genomic DNA was prepared by a modification of a method described elsewhere (Chakraborty et al., 2001). A single bacterial colony was inoculated in 10 mL of LB broth with 0.5% NaCl (w/v) and incubated in a rotary shaking incubator at 37 °C and 200 r.p.m. for 16 h. Overnight-grown bacterial cells were pelleted and resuspended in 2 mL of extraction buffer [100 mM Tris HCl (pH 8.0), 50 mM NaCl, 50 mM EDTA, 1% sodium dodecylsulphate (SDS)] to which lysozyme was added at a final concentration of 1 mg mL⁻¹ and incubated at 37 °C for 1 h. To this lysate, proteinase K was added at 100 μg mL⁻¹ concentration and incubated at 37 °C for 2 h and was extracted with a phenol/chloroform/isoamyl alcohol (25:24:1) mixture. To the aqueous supernatant, equal volume of isopropanol was added and pelleted. The pellet was washed with 70% ethanol, air-dried and suspended in TE buffer (10 mM Tris HCl, 1 mM EDTA).

ISSR primers

Eight ISSR primers including five di- and three trinucleotide repeats with a stretch of arbitrary nucleotides anchored at either 3' or 5' ends were used in the study (Table 1). All of these primers except (GA)₈T were designed in house at the Centre for DNA Fingerprinting and Diagnostics (CDFD) while the sequence of (GA)₈T was obtained from NAPS, University of British Columbia (UBC).

ISSR-PCR

ISSR-PCR amplification was carried out in a volume of 10 μL containing 1 × PCR buffer, 0.1 mM each of dCTP, dGTP, dTTP, dATP, 2.5 mM MgCl₂, 0.8 μM primer, 1.0 U of Taq DNA polymerase (MBI Fermentas) and 20 ng of template DNA. All the PCR amplifications were carried out on a PE 9700 (Perkin Elmer Corp) thermal cycler using the following reaction conditions: initial denaturation at 94 °C for 1 min and 35 cycles of 94 °C for 1 min, 50 °C for 30 s and 72 °C for 2 min followed by a final extension of 72 °C for 10 min. The PCR products were analyzed on a 2.5% MetaPhor agarose gel (0.5 × TAE buffer, 80 V). The gels were stained with ethidium bromide and documented using the Molecular Imager Gel Dc XR system (Bio-Rad Laboratories).

Detection of virulence genes

Using the sequence information available on the NCBI database, the primers were designed for the following Taq DNA polymerase (MBI Fermentas) and 20 ng of template DNA. All the PCR amplifications were carried out on a PE 9700 (Perkin Elmer Corp) thermal cycler using the following reaction conditions: initial denaturation at 94 °C for 3 min and 35 cycles of 94 °C for 1 min, 50 °C for 30 s and 72 °C for 2 min followed by a final extension of 72 °C for 10 min. The PCR products were analyzed on a 2.5% MetaPhor agarose gel (0.5 × TAE buffer, 80 V). The gels were stained with ethidium bromide and documented using the Molecular Imager Gel Dc XR system (Bio-Rad Laboratories).
virulence genes: TcpA, TcpP, TcpH, TcpB, TcpI, ToxT, which constitute the pathogenicity island CtxA, which codes for a cholera toxin subunit A and for the regulatory ToxR and ToxS genes. Primer pairs across the left and right junctions of the VPI – VPI5/VPI8 and VPI9/VPI10, respectively, and Cholera Toxin island (CTX) – rig/tlc3 and rtxA2/ctxB3, respectively, were also used for validating their absence (Boyd et al., 2000) (Table 1). PCR amplifications were performed in a reaction volume of 10 µL containing a final concentration of 1 × PCR buffer, 0.1 mM each of dCTP, dGTP, dTTP, dATP, 1.5 mM MgCl2, 1.0 µM primer, 1.0 U of Taq DNA polymerase (MBI Fermentas) and 20 ng of template DNA. All the PCR reactions were carried out in a PE 9700 (Perkin Elmer Corp) thermal cycler using the following reaction conditions: the initial denaturation of 94 °C for 3 min and 35 cycles of 94 °C for 30 s, annealing at respective temperatures for 30 s and 72 °C for 1 min followed by a final extension of 72 °C for 10 min. The PCR products were separated on a 2.0% metaphor agarose gel stained with ethidium bromide.

**ERIC-PCR**

ERIC-PCR amplification was performed as reported earlier (Rivera et al., 1995) in a reaction volume of 10 µL containing 1 × PCR buffer, 0.1 mM each of dCTP, dGTP, dTTP, dATP, 2.5 mM MgCl2, 1.0 µM primer, 1.0 U of Taq DNA polymerase (MBI Fermentas) and 20 ng of template DNA. All the PCR reactions were carried out on a PE 9700 (Perkin Elmer Corp) thermal cycler following the following reaction conditions: with the initial denaturation of 94 °C for 5 min and 35 cycles of 94 °C for 1 min, 52 °C for 1 min and 72 °C for 5 min followed by a final extension of 72 °C for 10 min. The PCR products were analyzed on a 1.5% metaphor agarose gel stained with ethidium bromide.

**Data analysis**

Specific ISSR-PCR and ERIC-PCR amplicons that were prominent and reproducible in two independent reactions were selected and scored for the presence as 1 or absence as 0. The sizing of the amplicons was carried out using the QUANTITY ONE software of the BioRad Gel Documentation system, against a size standard. Genetic distance was calculated from the data using WINDIST software of the WINBOOT package with the NTSYS format, and the genetic distance has been used for constructing dendrogram using the Neighbor-joining option of MEGA 3.2 software package (Kumar et al., 1994).

**Results and discussion**

ISSR-PCR profile brings about the variation in the frequency of SSR repeats among different strains of *V. cholerae*. In the *V. cholerae* genome, mono-, di- and trinucleotide repeats are the most abundant (Fig. 2). Among dinucleotide repeats, CG/GC (CG and GC being complementary, are considered as one group) repeats are the most abundant followed by AG/CT and, TGA/TCA is the most frequent followed by TGC/GCA among the tri-nucleotide repeats (Fig. 2). For the present study, a total of eight anchored ISSR primers that included five di- and three tri-nucleotide repeat containing primers (Table 1) were chosen based on the SSR abundance allowing for a better coverage of the whole genome and scorability of the marker profile on a metaphor agarose gel electrophoresis. For considering a marker as polymorphic, the absence of an amplified product in at least one strain in a pairwise strain comparison was treated as the criterion.

All eight primers used in this study produced a total of 221 bands ranging from 190 to 2500 bp and all of them were found to be polymorphic, with 100% reproducibility (Table 2). A representative ISSR-PCR profile generated by primer (GA)6C (Fig. 3) and its comparison with the ERIC-PCR profile (Fig. 4) revealed the level of informativeness of ISSR markers. The 5’-anchored dinucleotide primer (GA)6T generated the least number of 16 markers as against the highest number of 37 markers generated by 3’ anchored

---

Fig. 2. (a) Whole genome profile of microsatellite abundance in *Vibrio cholerae*: Motifs ranging from mono- to hexa-nucleotides repeating a minimum of three times were considered. The numbers at the top of the bars are the actual number of the respective SSR type present in the genome. (b) Whole-genome profile of dinucleotide repeat abundance in *Vibrio cholerae* (complementary repeat motifs are grouped into one class, for e.g., CA/GT etc). (c) Whole-genome profile of tri-nucleotide repeat abundance in *V. cholerae* (the three different possible tri-nucleotide repeat units and their complementary motifs are grouped into one class for e.g., TGA/TCA, GAT/ATC, ATG/CAT etc).
trinucleotide primer TA(CAG)₄. Among O1 El Tor strains, the primers (GA)₈C and (GA)₈T generated the highest and lowest levels of polymorphism of 100% and 75%, respectively, and among all the O139 strains, the primers (ATG)₄GA, (CA)₇C, (GA)₈C, (GA)₈T, TA(CAG)₄ and T(GA)₈ generated 100% polymorphism and the lowest

### Table 2. Amplicon profile generated by the eight ISSR primers

<table>
<thead>
<tr>
<th>Primer code</th>
<th>Amplicon size range (bp)</th>
<th>No. of markers</th>
<th>No. of polymorphic markers</th>
<th>O1</th>
<th>O139</th>
<th>O1 and non-O1</th>
<th>O139 and non-O139</th>
<th>O1 and non-O1 non-O139</th>
<th>O139 and non-O1 non-O139</th>
</tr>
</thead>
<tbody>
<tr>
<td>(ATG)₄GA</td>
<td>250–1700</td>
<td>25</td>
<td>25</td>
<td>1700</td>
<td>–</td>
<td>–</td>
<td>250, 400</td>
<td>–</td>
<td>950</td>
</tr>
<tr>
<td>(CA)₇C</td>
<td>230–2500</td>
<td>29</td>
<td>29</td>
<td>–</td>
<td>–</td>
<td>360</td>
<td>250, 300, 340</td>
<td>470</td>
<td>230</td>
</tr>
<tr>
<td>CGA)₇</td>
<td>190–1300</td>
<td>32</td>
<td>32</td>
<td>–</td>
<td>–</td>
<td>580</td>
<td>350, 520</td>
<td>980</td>
<td>–</td>
</tr>
<tr>
<td>(GA)₈C</td>
<td>300–2000</td>
<td>27</td>
<td>27</td>
<td>–</td>
<td>–</td>
<td>550</td>
<td>470</td>
<td>–</td>
<td>230</td>
</tr>
<tr>
<td>(GA)₈T</td>
<td>420–2400</td>
<td>16</td>
<td>16</td>
<td>900</td>
<td>–</td>
<td>1000, 2400</td>
<td>500, 1250</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GC(GCC)₄</td>
<td>200–1900</td>
<td>26</td>
<td>26</td>
<td>–</td>
<td>–</td>
<td>1300, 1900</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>TA(CAG)₄</td>
<td>350–2000</td>
<td>37</td>
<td>37</td>
<td>–</td>
<td>–</td>
<td>450, 520</td>
<td>380, 500</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>T(GA)₈</td>
<td>190–1900</td>
<td>29</td>
<td>29</td>
<td>–</td>
<td>800</td>
<td>590</td>
<td>550, 620, 750, 950</td>
<td>700</td>
<td>–</td>
</tr>
</tbody>
</table>

**Fig. 3.** ISSR-PCR profile for non-O1/non-O139, O1 and O139 strains generated by (GA)₈C primer run on 2.5% agarose gel. M1 – 100 bp size marker (MBI Fermentas), M2 – 1 kb size marker (MBI Fermentas). Arrows indicate marker sizes.

**Fig. 4.** ERIC-PCR profile of the same set of non-O1/non-O139, O1 and O139 strains used in Fig. 3, run on 1.5% agarose gel. M1 – 100 bp size marker (MBI Fermentas), M2 – 1 kb size marker (MBI Fermentas). Arrows indicate marker sizes.
polymorphism of 80% was observed with primer GC(GCC)₄. The number of polymorphic markers generated and the percentage of polymorphism with each of the primers for toxinogenic O1 El Tor and O139 strains decreased drastically, with the exclusion of nontoxinogenic O1 El Tor strain G-5 and nontoxinogenic O139 strains B-3 and B-27 (Table 3).

Phylogenetic analysis

The dendrogram generated based on the cluster analysis showed the grouping of O1 El Tor strains with O139 strains in a single clade indicating the genetic similarity between these two sets of strains and thus corroborating the hypothesis of evolution of O139 strains from an ancestral O1 El Tor strain with the exchange of O antigen by horizontal gene transfer (Calia et al., 1994; Stroeher et al., 1995; Faruque et al., 2003). Furthermore, the finer subclustering within the serogroups reflects on the clonal nature of the strains. For example, the strains N-33 and N-35, N-36 and N-37 grouped into one cluster suggesting their clonal origin. The same was the case with N-1, N-8, N-7 and N-11 strains while the numerous nontoxinogenic singletons suggest their diverse sources of origin (Thompson et al., 2005). In the present study however, one O1 El Tor strain (G-5) and two O139 strains (B-3 and B-27) were assigned to the diverse non-O1/non-O139 group, which are nontoxinogenic in nature. Screening for the presence of VPI and CTX islands across junctions of VPI (VPI5/VPI8; VPI9/VPI10) and CTX (rig/tlc3; rtxA2/ctxB3) were also used. Toxinogenic O1 El Tor strain G-2 and toxinogenic O139 strain B-36 are positive controls. Arrows indicate the amplicon sizes.

Comparison of ERIC-PCR and ISSR-PCR profiles

The dendrogram generated using the ERIC-PCR data was similar to that of ISSR-PCR data, with respect to the overall clustering of toxinogenic and nontoxinogenic strains of V. cholerae (Fig. 6). However, in ISSR-PCR, within the clade of toxinogenic strains, a distinct subclustering of O139 and O1 El Tor strains was observed while no such resolution of clusters was observed in the ERIC-PCR-based dendrogram.
The same was the case with non-O1/non-O139 isolates, the grouping of which is clearly distinct in the ISSR-PCR dendrogram. Thus, ISSR-PCR is technically simple, robust, high-throughput, universally applicable and cost-effective and could be exploited as an efficient tool in the phylogenetic and molecular epidemiological studies of any prokaryotic genome. Moreover, specific markers for specific pathogens could be identified, standardized and developed as a ready-to-use detection kit.

**Acknowledgement**

V.S. is a recipient of Research Fellowship by Council for Scientific and Industrial Research, Government of India.

**Authors contribution**

A.R. and V.S. contributed equally to this work.

**References**


